**Thinking like a Data Scientist**

**In this Assignment, you will demonstrate your understanding of the data science methodology by applying it to a given problem. Pick one of the following topics to apply the data science methodology to:**

1. **Fake News Detection**
2. **Sentiment Analysis (Detection)**
3. **Emails (Detecting spam emails)**
4. **Hospitals (Providing better health care)**
5. **Credit Cards (Best spender)**
6. **Credit Cards Fraud Detection**
7. **Food Recipes (Best seller meal)**
8. **MySejahtera (Detecting crowded areas, new clusters)**
9. **Loan Payment**
10. **Car Price**
11. **Stock Market**
12. **Any other topic you are interesting in.**

**You will have to play the role of a data scientist, you have to come up with a problem that is more specific but related to one of these topics.**

# **Trending Food**

## Business understanding

The main problem

* Restaurant or grocer need to create new **food** in line with current trends and food demand.

(BTS meals, Burger Nasi Lemak, Spicy Fried Chicken, Bubble tea)

Objective

* To give Restaurant or grocer suggestions of new trendy and demanded food.

## Analytics Approach

NO

YES

Unrelated

NO

YES

Unrelated

Trending Food

Posting in SocMed

About Food

Trending Food

Food Photo

## Data Requirement

## Listing down the required data content, the format of the data.

## Save data in BLOB format.

Source:

* Instagram
* #food
* #foodie
* #foodporn.
* #instafood.
* #foodphotography.
* #yummy.
* #foodstagram.
* #foodblogger.

## Data Collection

Set row and column and determine whether data collected suit with finding trending food.

## Data Understanding

Understanding the data and studying the data set.

Improving data quality by removing missing value and invalid and misleading data.

## Data Preparation

Understanding rows and columns, manipulate the data.

Set data to 70% training data to train system, 30% testing data to test system accuracy.

## Modelling

Developing model based on the chosen machine learning algorithm and implementing classification algorithm.

## Evaluation

Evaluating results of machine learning. Comparing the results with expected results, improving by tuning the parameters.

## Deployment

Deploying model on the real, new data.

## Feedback

Receiving feedback from user.